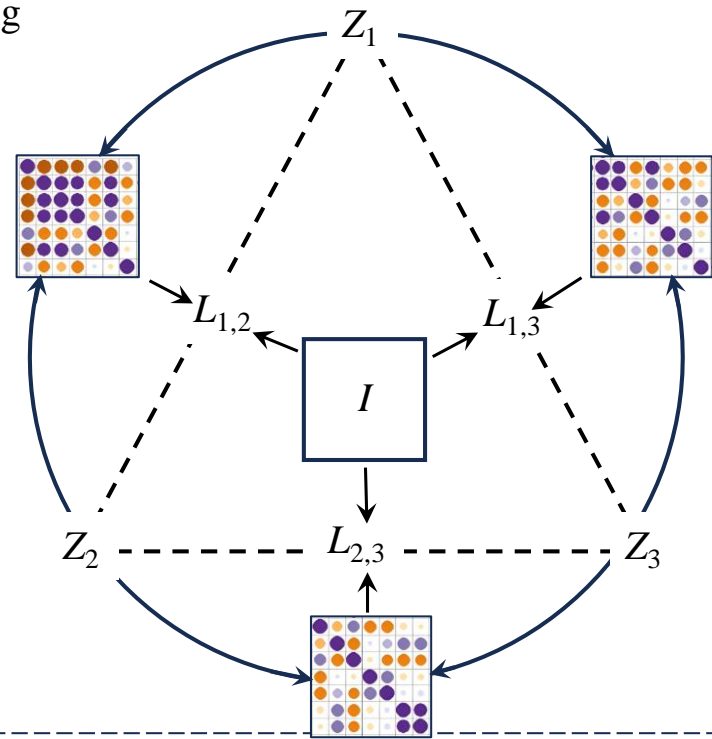
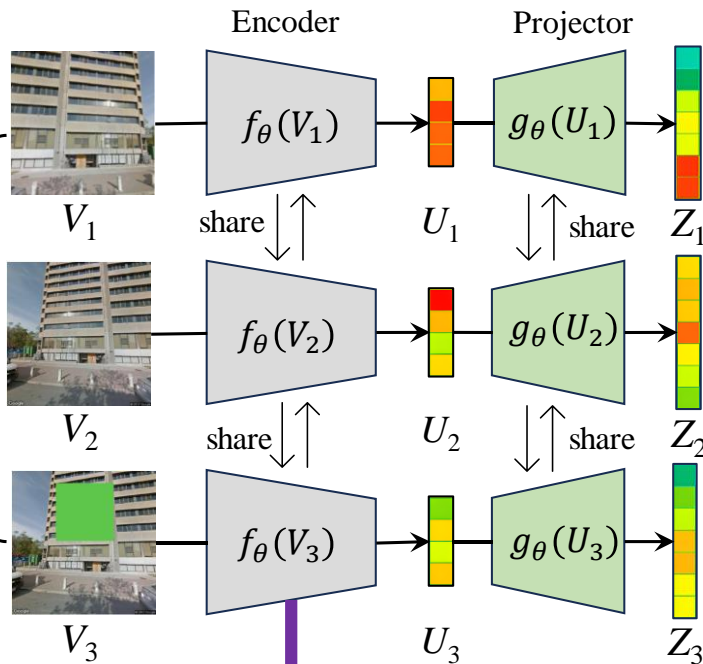


Phase 1: Self-supervised pre-training

Street view images
without label



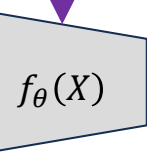
Tri-branch
mutually-
exclusive
augmentation



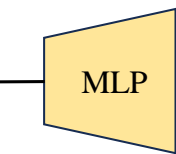
Street view
images with label



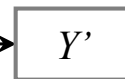
(X,Y)



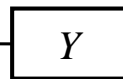
Freeze
Encoder



Classification
Head



L_{CE}



Phase 2: Supervised fine tuning